The Breast 37 (2018) 80–88

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Metastasis-free interval in breast cancer patients: Thirty-year trends and time dependency of prognostic factors. A retrospective analysis based on a single institution experience

A. Houzé de l'Aulnoit [a, 1], B. Rogoz [a, *], C. Pinçon [b, 2], D. Houzé de l'Aulnoit [a]

[a] Department of Obstetrics and Gynaecology, Saint Vincent de Paul Hospital, Catholic University of Lille, Lille, France
[b] Laboratory of Biomathematics, University of Lille, CHU Lille, EA 2694, Lille, France

ARTICLE INFO

Article history:
Received 21 May 2017
Received in revised form
16 October 2017
Accepted 18 October 2017

Keywords:
Breast cancer
Metastasis-free interval
Cox model
Prognostic factors

ABSTRACT

Introduction: Breast cancer remains the leading cause of cancer death in French women in spite of continuously improving management. The objectives of this study were to analyse trends in the metastasis-free interval over the past 30 years and to identify the prognostic factors of survival, while accounting for time dependency.

Methods: A total of 1613 patients diagnosed with invasive non-metastatic breast cancer at Saint Vincent de Paul Hospital, Lille, France between 1977 and 2013, were followed for outcome (metastasis-free interval). Cohort entry time delay, a continuous temporal covariate, was defined to assess improvement of outcome. Data were analysed using the Cox proportional hazards model and presented as hazard ratio (HR).

Results: Metastatic disease developed during follow-up in 446 (27.6%) patients. Cohort entry time delay exhibited strong independent prognostic value while accounting for multiple prognostic factors including: tumour size (HR = 1.62, 95 %CI 1.37–1.91); rapid tumour growth (HR = 1.59, 95%CI 1.17–2.16); lymph node ratio (HR = 2.29, 95%CI 1.97–2.66); histological grade (grade 2 was significant only during the first 10 years after diagnosis, grade 3 and progesterone receptor status only during the first 5 years after diagnosis); and oestrogen receptor status (significant only during the first 8 years (HR = 0.75, 95%CI 0.58–0.96)).

Conclusion: The current study showed an improvement in the prognosis of breast cancer patients over the past 30 years and pointed to the importance of evaluating covariates with time-varying effects.

1. Introduction

Breast cancer is still the leading cause of cancer death in French women, even though mortality has significantly declined over the past 30 years [1,2]. Substantial efforts have been made to identify the prognostic factors that are associated with better or worse outcome [3] and the management of breast cancer patients has known continuous improvement [4–8]. Recent statistics made in the United States, reported 98% 5-year overall survival for local breast cancer, with 84% for regional breast cancer and 23% for metastatic disease [9].

Survival analysis, or time-to-event data analysis, is widely used in oncology, the studied event being cancer recurrence, metastasis or death. Prognosis can be assessed in terms of overall survival, progression-free survival or metastasis-free interval (MFI). The effects of prognostic factors can change over time and better evaluations are expected with the use of adapted statistic models. Little work has been devoted to the long-term impact of improved management practices on survival in breast cancer patients, and is often reported with unclear statistical methodology [10]. The first objective of this study was to use a continuous temporal covariate transcribing the global therapeutic effect on metastasis-free survival to analyse the outcome of breast cancer patients over the past 30 years. The second objective of the study was to identify the

* Corresponding author. 3 Boulevard Montebello, Bâtiment A4, Apt 432, 59000, Lille, France.
E-mail addresses: Houze-De-L-Aulnoit.Agathe@ghicl.net (A. Houzé de l'Aulnoit), rogoz.bianca@ghicl.net (B. Rogoz).
[1] Boulevard de Belfort, 59000, Lille, France.
[2] Laboratory of Biomathematics, University Lille 2, 3, rue du Professeur Laguesse, BP 83, 59006, Lille, France.

https://doi.org/10.1016/j.breast.2017.10.008
0960-9776/

prognostic factors of metastasis-free survival, while accounting for time dependency.

2. Patients and methods

Our institutional review board, the ethics committee at the Catholic Institute of Lille, approved this retrospective study.

2.1. Patients

This study included 1727 women who were treated for invasive breast cancer between January 1977 and December 2013 at the Saint Vincent de Paul Hospital, Lille, France. All patients were staged according to the international classifications applicable at the time of diagnosis and given treatments in compliance with the current national protocols.

2.2. Data collection

Data were retrieved from our department's database that was created with the approval of the National Commission on Informatics and Liberty in 1990. At that time, data (on paper charts) from patients treated for invasive breast cancer between January 1977 and 1990 were registered in the database retrospectively to facilitate future studies. Since 1990, data have been registered at the time of diagnosis. Tumour characteristics and staging were recorded in the database according to the international classifications applicable at the time of diagnosis.

No modification of initial data was made for the purpose of this study and missing items were not replaced. We collected information related to the primary breast cancer: age at diagnosis, year of diagnosis, tumour and node clinical stage, tumour localisation, histological characteristics (tumour size, grade, lymph nodes, oestrogen receptor (ER) status, progesterone receptor (PR) status), and context of rapid tumour growth using the PEV classification proposed by P. Denoix at the Gustave-Roussy Institute in France in 1970 [10]. Rarely used outside of France, PEV (for potential evolution) represents a subjective evaluation of the evolution of breast cancer based on two criteria – tumour growth from the time of discovery by the patient to the first medical examination (maximum 6 months) and presence of inflammatory signs – as follows: PEV0 = absence of a rapid evolution of the tumour, with a normal period of tumour growth (doubling time of more than 110 days); PEV1 = rapid evolution of the tumour, with a doubling time of less than 110 days; PEV2 = local inflammation; PEV3 = inflammatory carcinoma of the whole breast [10]. Tumour size was systematically available only after 1990 and hormone receptor status after 1996. HER2 status (recorded after 2005) was not studied because of the amount of missing data (82.8%).

2.3. Study design

We studied the MFI defined as the time between the diagnosis of primary breast cancer and the diagnosis of metastatic disease. For patients who did not develop metastatic disease, the endpoint of this interval was considered as the date of the last contact with the patient. Patients who died from another cause, patients who had a previous or concomitant malignant disease, and patients who developed a second, contralateral breast cancer were censored. Patients who developed only local or regional recurrence were retained for analysis, as they were considered non-metastatic, and continued to be followed for metastasis. We excluded the patients with initial or de novo metastatic breast cancer (MFI < 3 months) (n = 114) [11,12] (Fig. 1). A total of 1613 patients were thus eligible for analysis.

In order to evaluate the improvement in the outcome of breast cancer over the past 30 years, we defined a continuous temporal variable: “cohort entry time delay”. Measured in years, the cohort entry time delay was defined as the time between the beginning of the study (January 1977) and the year of diagnosis of the patient's invasive breast cancer. Earlier work has also used this variable to study the outcome of metastatic disease [13].

2.4. Statistical analysis

2.4.1. Descriptive analysis

Missing data were treated by a single imputation approach. When the missing data represented more than 50%, the covariate was excluded from the model. The median follow-up and its confidence interval were estimated by bootstrapping. Continuous variables are presented as mean value and standard deviation in case of normal distribution, or as median value [25th-75th percentile]. Categorical variables are presented as number of patients and percentages.

2.4.2. Univariate analysis

For categorical variables, we used the log-rank test and the Kaplan-Meier method. Regrouping was performed whenever no statistical difference was observed among different groups or less than 5% of the patients were allocated to a certain group. For continuous variables, we used the Wald test and tested their log-linearity. Indeed, the coefficient β of a covariate must be constant and must not change with the decrease or increase of the covariate value. If this hypothesis was not verified, the covariate was transformed.

For all variables, we then tested the assumption of proportional hazards (PH) using a Cox univariate model, in which the hazard ratio (HR) is assumed constant over time. This means that the prognostic value of the studied variable will not change with time.

We identified the prognostic factors with a 0.20 level of significance to the mode and then added all the possible interactions between these explanatory variables.

2.4.3. Multivariate analysis

Imposing first the covariate “cohort entry time delay” to the model, we performed the selection of the pertinent covariates both manually and automatically, using the stepwise procedure according to the Akaike information criterion method. The assumption of PH was studied using graphics and tests based on scaled Schoenfeld residuals and cumulative martingale residuals [14]. As the assumption was not checked for several variables, suggesting that the HR is not constant over time, we next fitted an Extended Cox model allowing covariates to have time-varying effects [15–18]. Several options are available in order to account for time-dependency [14,19]:

- by associating them a covariate-by-time interaction,
- as binary variables, stratifying time in two intervals (before and after a certain time which maximizes the log-likelihood [17,20]),
- or as multiple binary classes corresponding to successive temporal intervals [15] according to Schoenfeld residues graphs.

The accuracy of the model was measured using Harrell's concordance index *c*, O'Quigley and Royston's indices R^2 and the calibration slope, after previous bootstrapping (n = 100) [16,21]. We estimated the effect of each variable by using the HR per unit and the percentage of the risk variation [16,18].

All statistical analyses were performed using the R statistical package. A p-value < 0.05 was considered statistically significant.